



QCon 全球软件开发大会
INTERNATIONAL SOFTWARE
DEVELOPMENT CONFERENCE

BEIJING 2017

高性能高可用机票实时搜索系统

去哪儿网 梁启康

议题

系统诉求

海量数据

设计思路

搜索框架

报价引擎

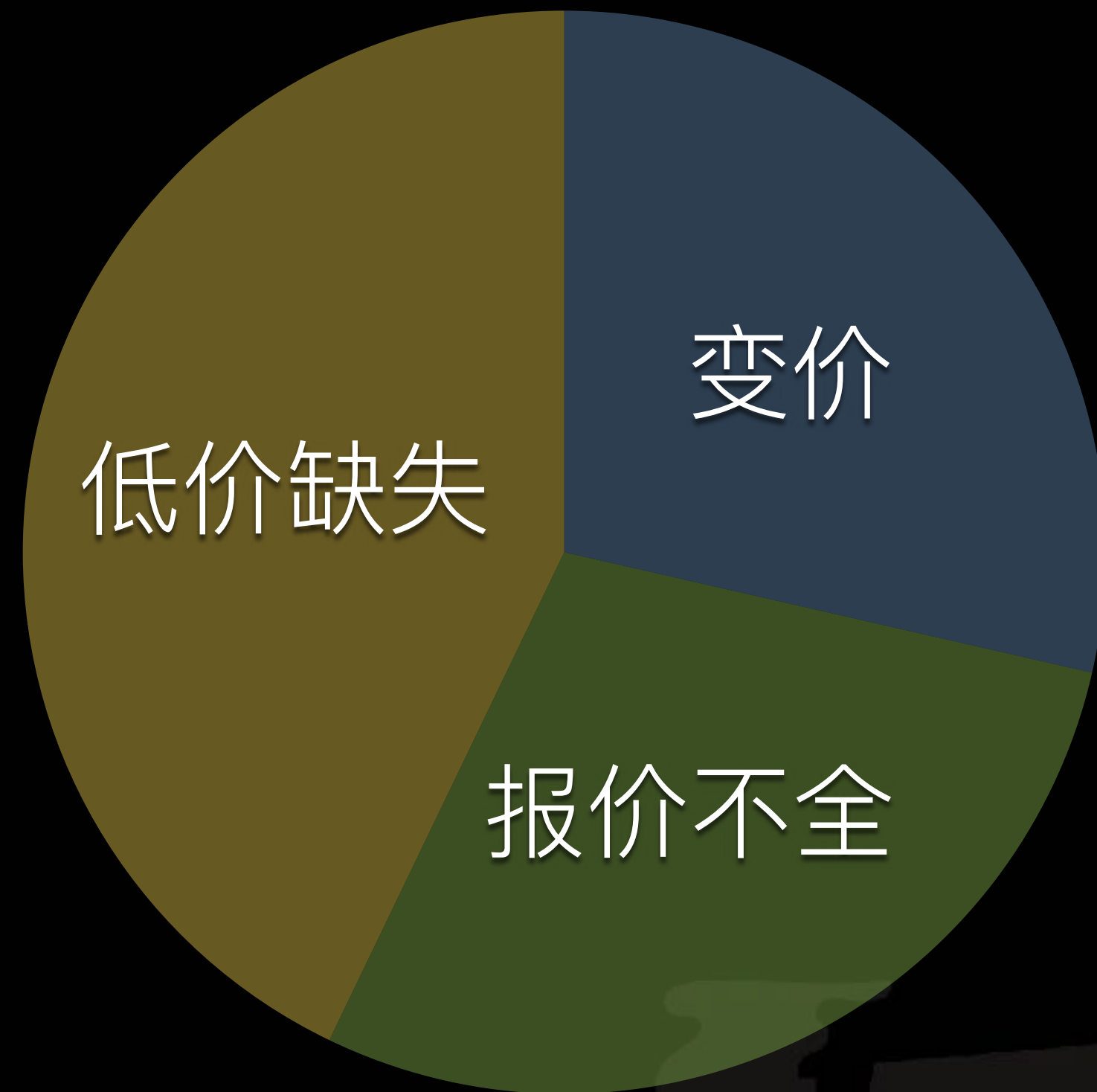
待解问题

系统诉求

- 全网价**最低**
- 航线报价**最全**
- 实时性**最好**
- 产品**最**丰富
- 预定**最**流畅



面临的问题



航班舱位时刻变动

供应商规则调整密集

运价规则变化繁多

航司政策各有不同

GDS数据成本不菲

供应商的office权限不一致

海量数据

- 规则量

- 供应商规则量2亿+
- 运价规则1亿+

- 报价量

- 国内国际航线28万 * 航班数 * 产品个数 * 供应商个数 * 180天 = 1500亿 +

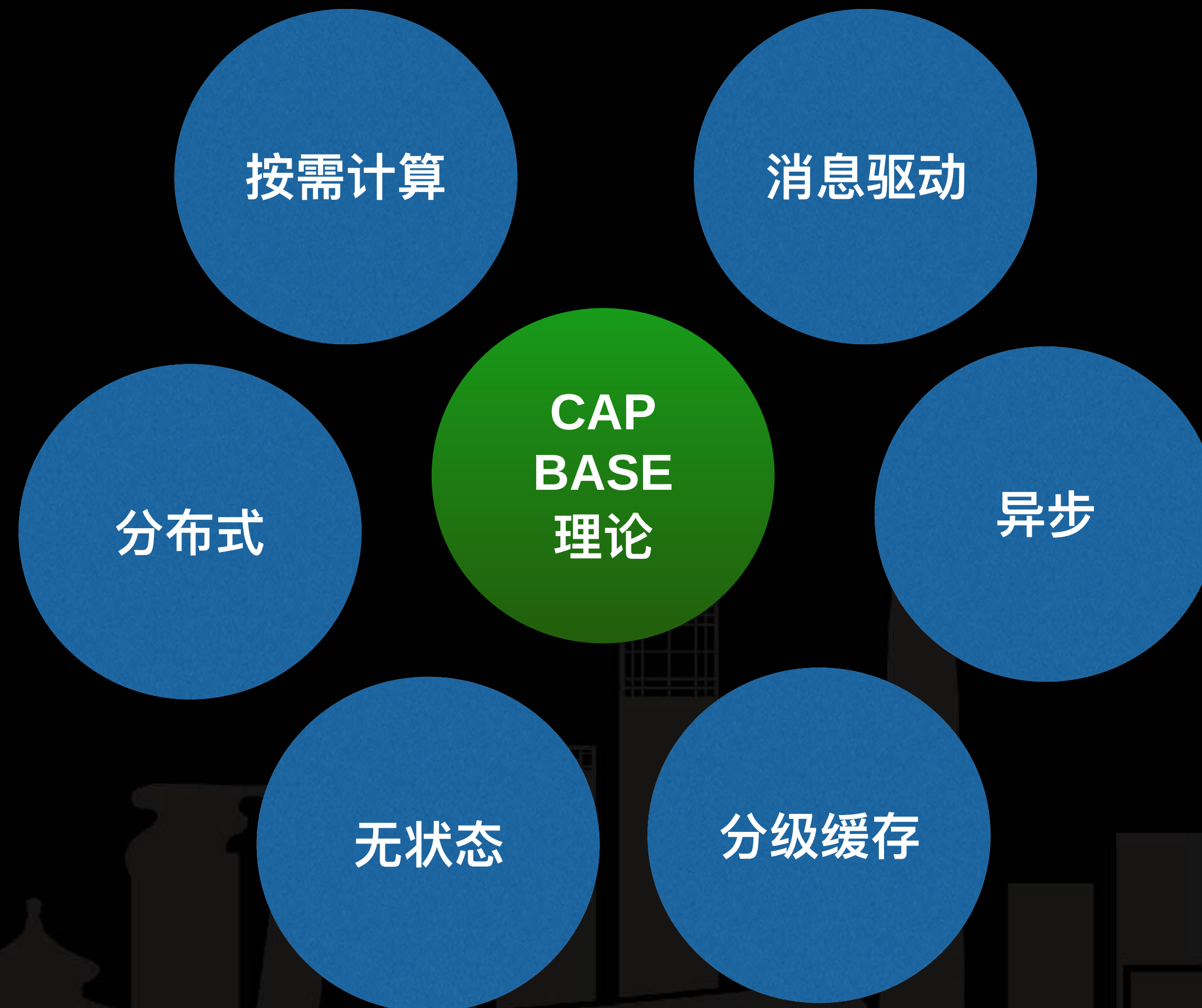
- 搜索量

- 航线搜索+航班搜索: 3k+qps
- 每秒计算产品数: 搜索qps * 航班数 * 供应商个数 * 产品个数 = 1500万 +

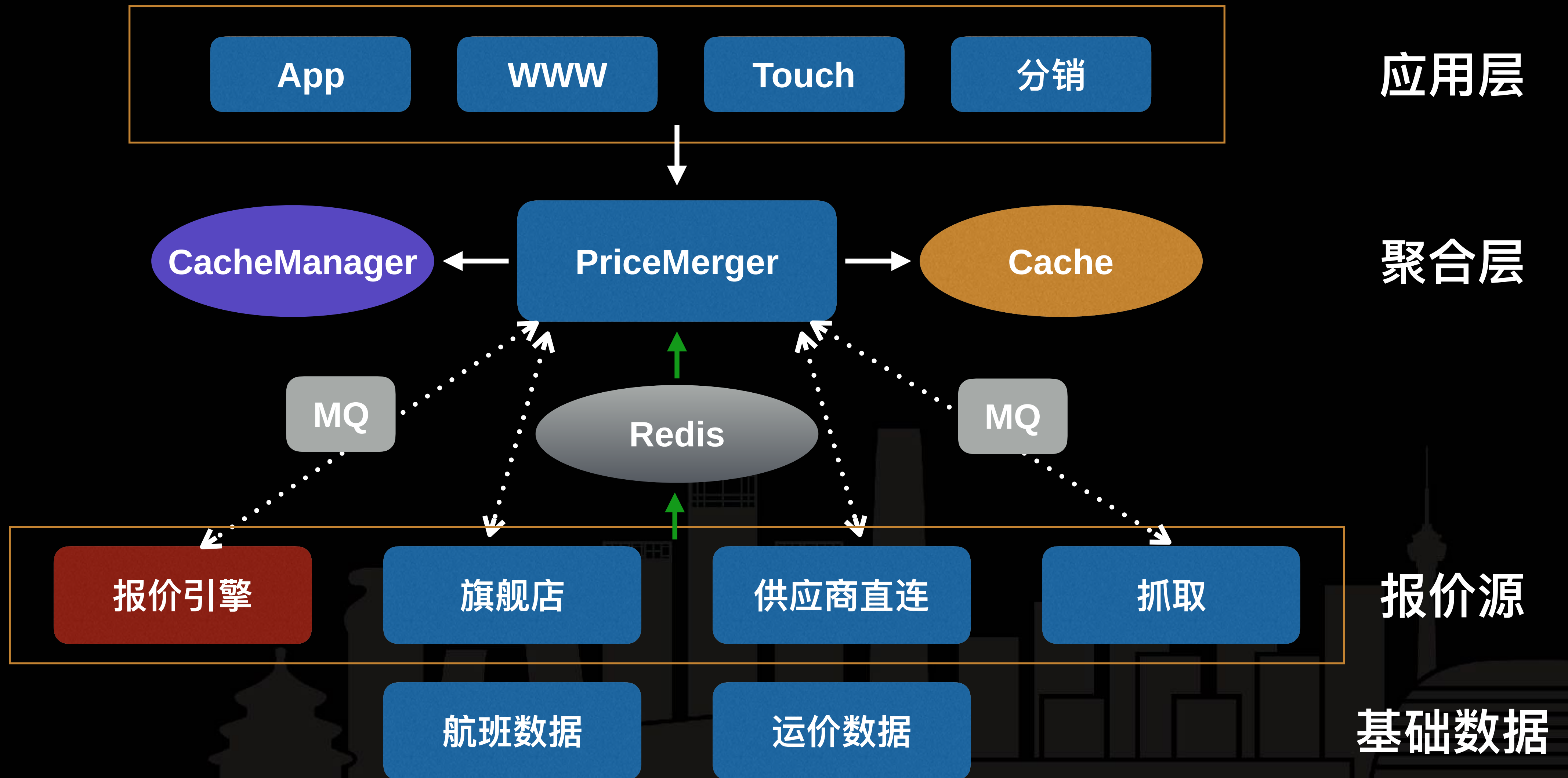
设计思路

- 最理想的方式
 - 所有的报价都实时计算，填充好
 - 一个巨大的哈希表
 - 响应所有渠道价格变动进行计算
 - 快
 - 0 变价

设计思路



搜索框架



搜索框架

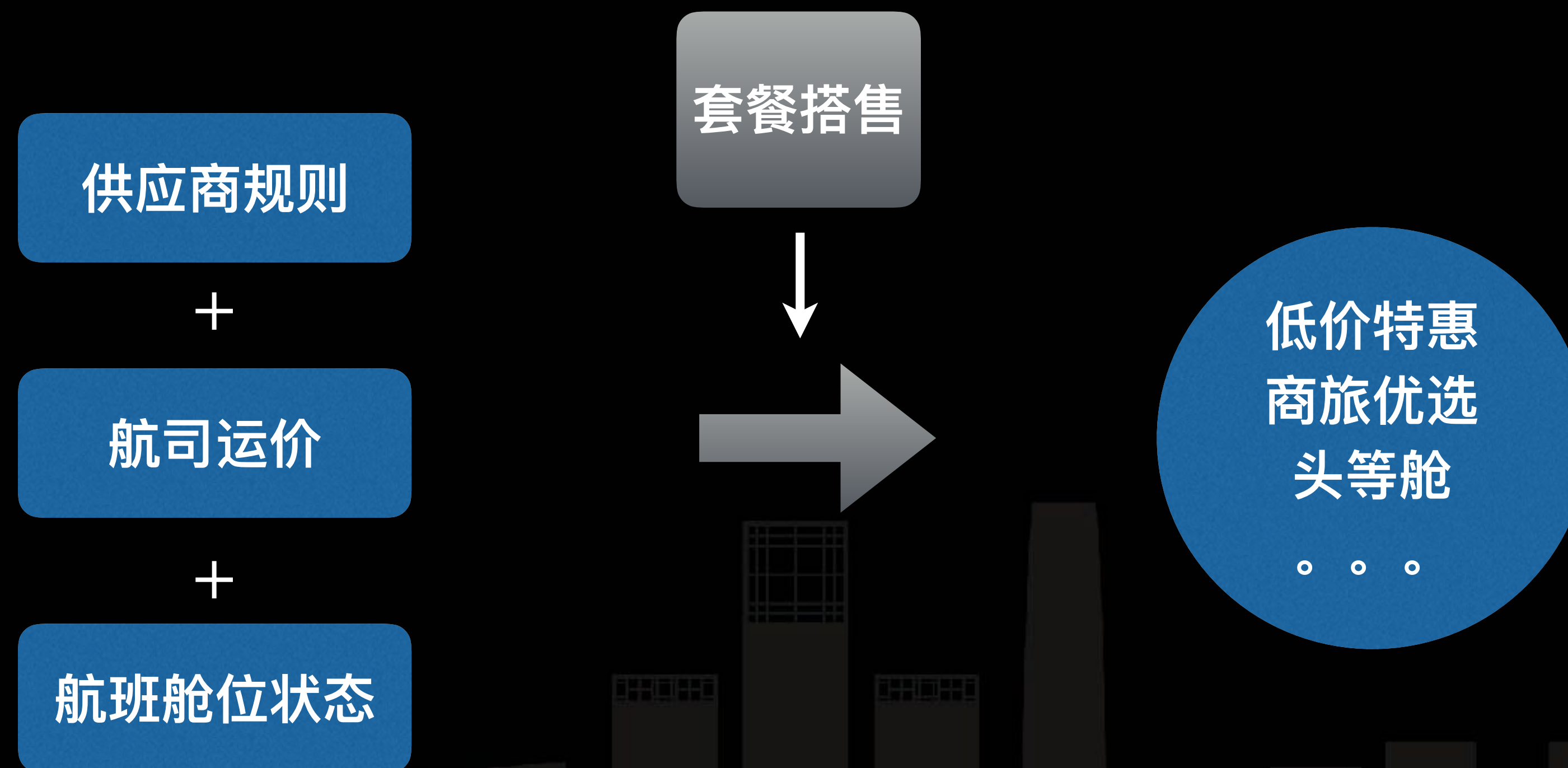
- 报价数量多，内存问题
- 报价源多，不稳定
- 响应时间要求苛刻
- 搜索条件有热点
- 缓存报价不准

遇到问题

解决方案

- 尽量用原生数据类型，编码
- 回数比例，异步搜索更新
- 分批回数
- 一致性哈希，负载均衡
- 供应商粒度，主动更新

报价引擎



报价引擎—供应商规则

☐ 申请产品☒ 包机切位产品☐ 预付产品

政策号:	XLB170322182348000102		
AV校验舱位表:	如果想对适用舱位进行AV校验,请 录入 ,舱位开放,该政策才会生效		
*航空公司:	CA	一条政策只能对应一个航空公司,只支持航空公司二字码输入,如:CA	
政策代码:	123		
禁售时间:	0	只能输入正整数,输入范围为60~120 (含60和120),不在范围内的数值默认为120	
*销售有效期:	从 2017-03-09	至 2017-04-29	☐ 旅行有效期相同 当天截止销售的政策,18点之后不显示
*旅行有效期:	从 2017-03-09	至 2017-04-29	输入格式:YYYY-MM-DD,如:2010-01-02
最晚提前出票时限:	2	含当天,只能输入正整数,且必须大于等于0。例如:最晚出票时限为2,用户至少需要提前2天预订机票。	
*班期限制:	☒ 1 ☐ 2 ☒ 3 ☐ 4 ☒ 5 ☐ 6 ☐ 7 周一至周日,周日用7表示		
*是否提供行程单	☒ 行程单 ☐ 行程单和差额发票 ☐ 仅提供发票		
是否有常旅客积分:	☒ 是 ☐ 否		
证件类型:	☒ 所有证件 ☐ 只支持身份证		
最大购买年龄:		0-100整数,允许购买机票的最大年龄;为空表示无年龄上限限制	
最小购买年龄:		0-100整数,允许购买机票的最小年龄;为空表示无年龄下限限制	
*起飞机场:	PEK	只支持一个起飞机场,输入只支持机场三字码,如PEK	
*到达机场:	CTU	仅普通政策支持多个到达机场,多个机场用","分隔,最多支持25个机场。特价、特殊政策只支持一个到达机场。	
*航班号:	☐ 全部航班 ☐ 不适用航班 ☒ 适用航班 CA1011,CA1345,CA2305,CA2306,CA2408 说明:支持录入多个航班与航班范围,范围之间用“-”链接,用","分隔。单个航班需要在航班范围前录入。 例如:MU8072,MU2312-MU5431		

报价引擎

• 背景

- 供应商定价规则、航班舱位、运价，决定底价
- 供应商数千，不少供应商上千万条规则
- 定价规则按供应商分库

• 问题

- 供应商规则更新频繁
- 单次搜索要计算所有供应商规则
- DB读写压力大
- 舱位状态变化频繁
- 热点航线问题
- 计算量巨大

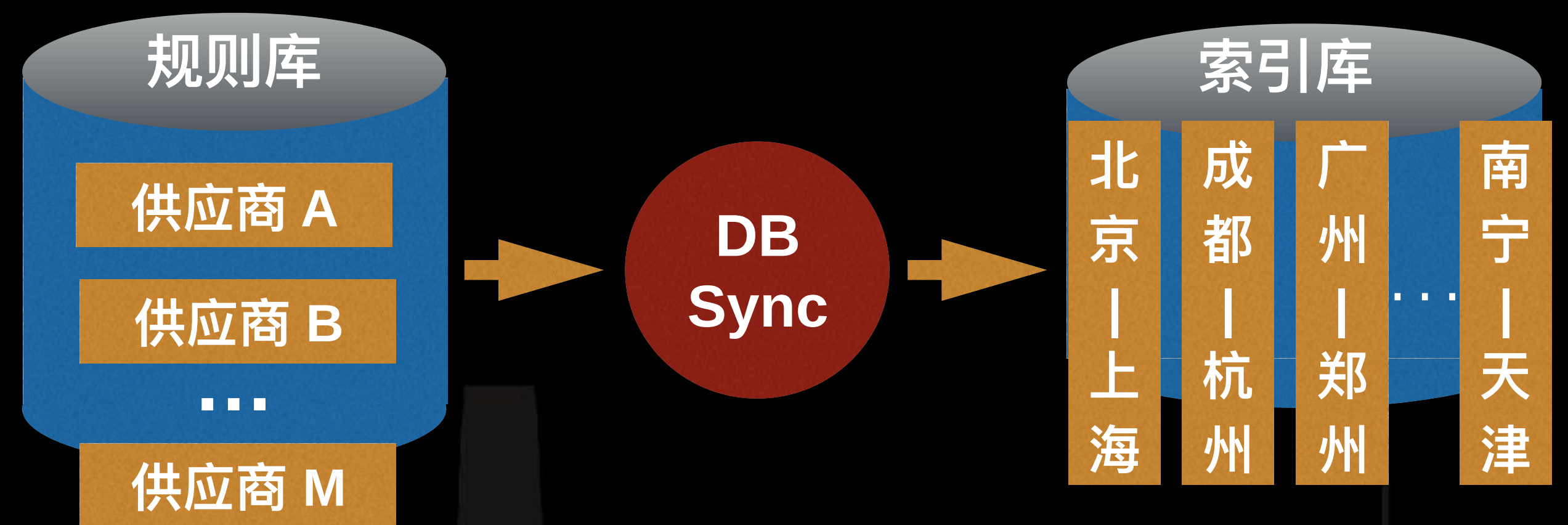
报价引擎

搜索引擎核心

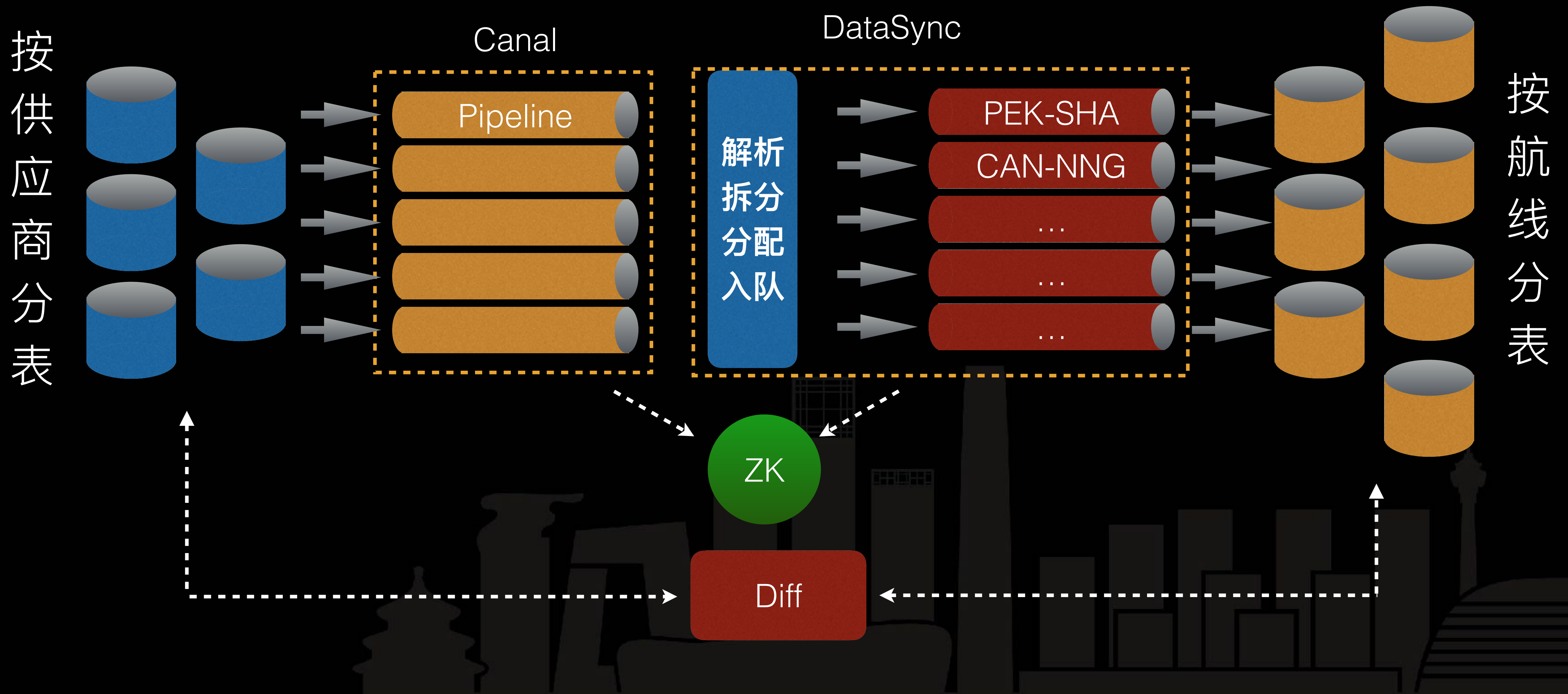
索引

报价引擎 — 组织索引

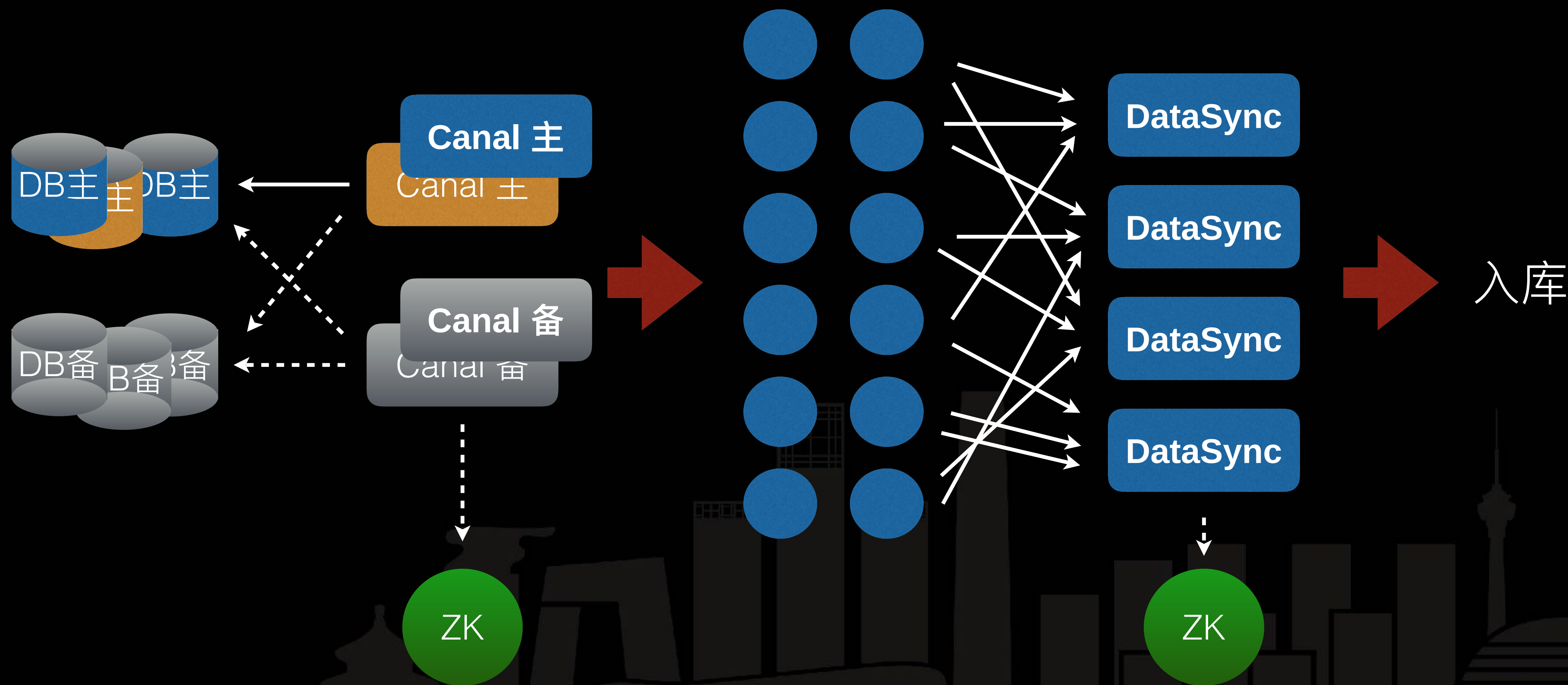
- 规则库 => 索引库
- 规则库写入量大，集群峰值达20K TPS
- 要求同步延迟很低，不超过60s
- 保持顺序一致性，如果先删后插变成先插后删，数据会不一致
- 数据最终一致
- 系统高可用



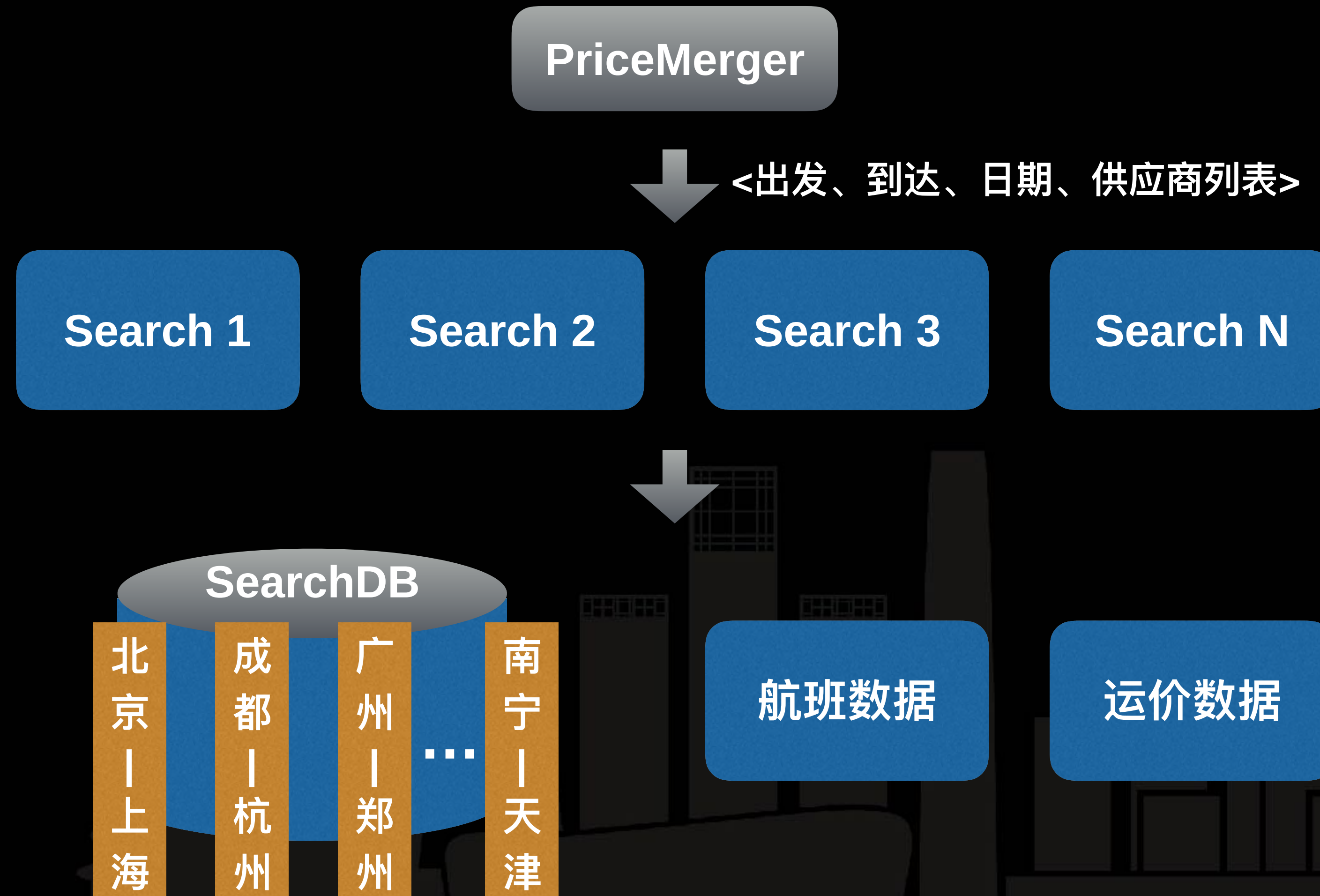
报价引擎 — 数据同步



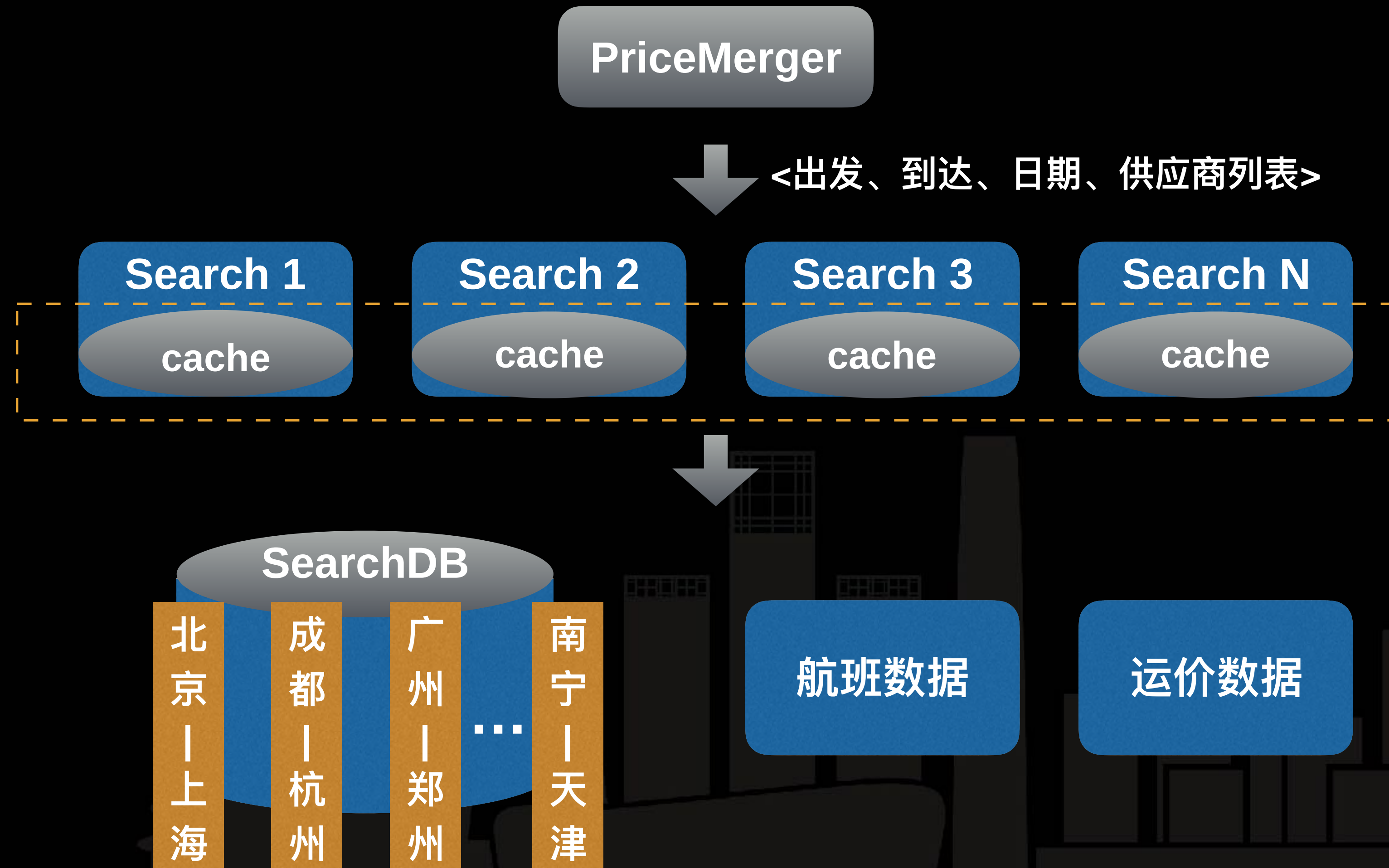
报价引擎 — 同步系统高可用



报价引擎 — 架构

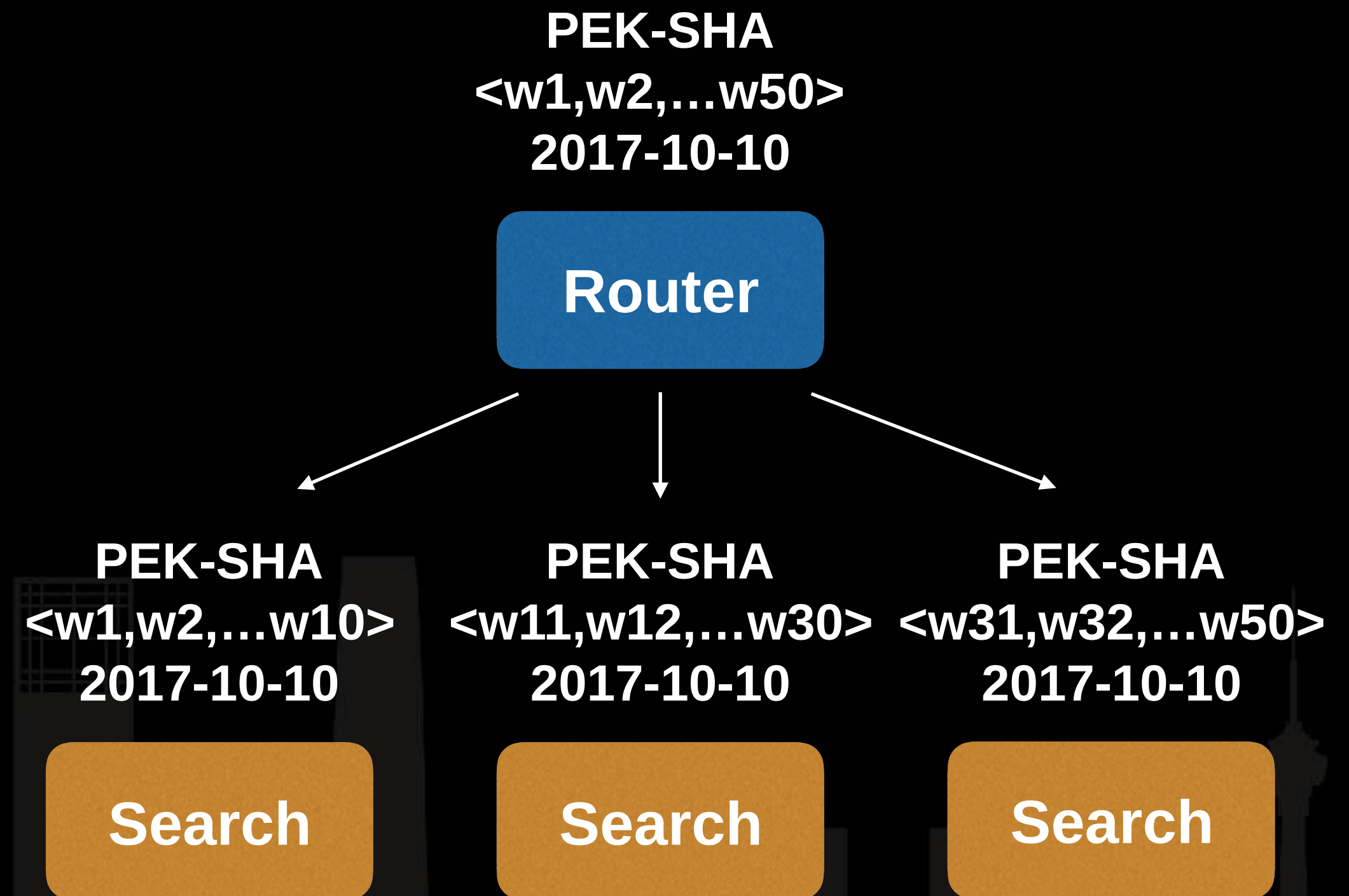


报价引擎 — 架构

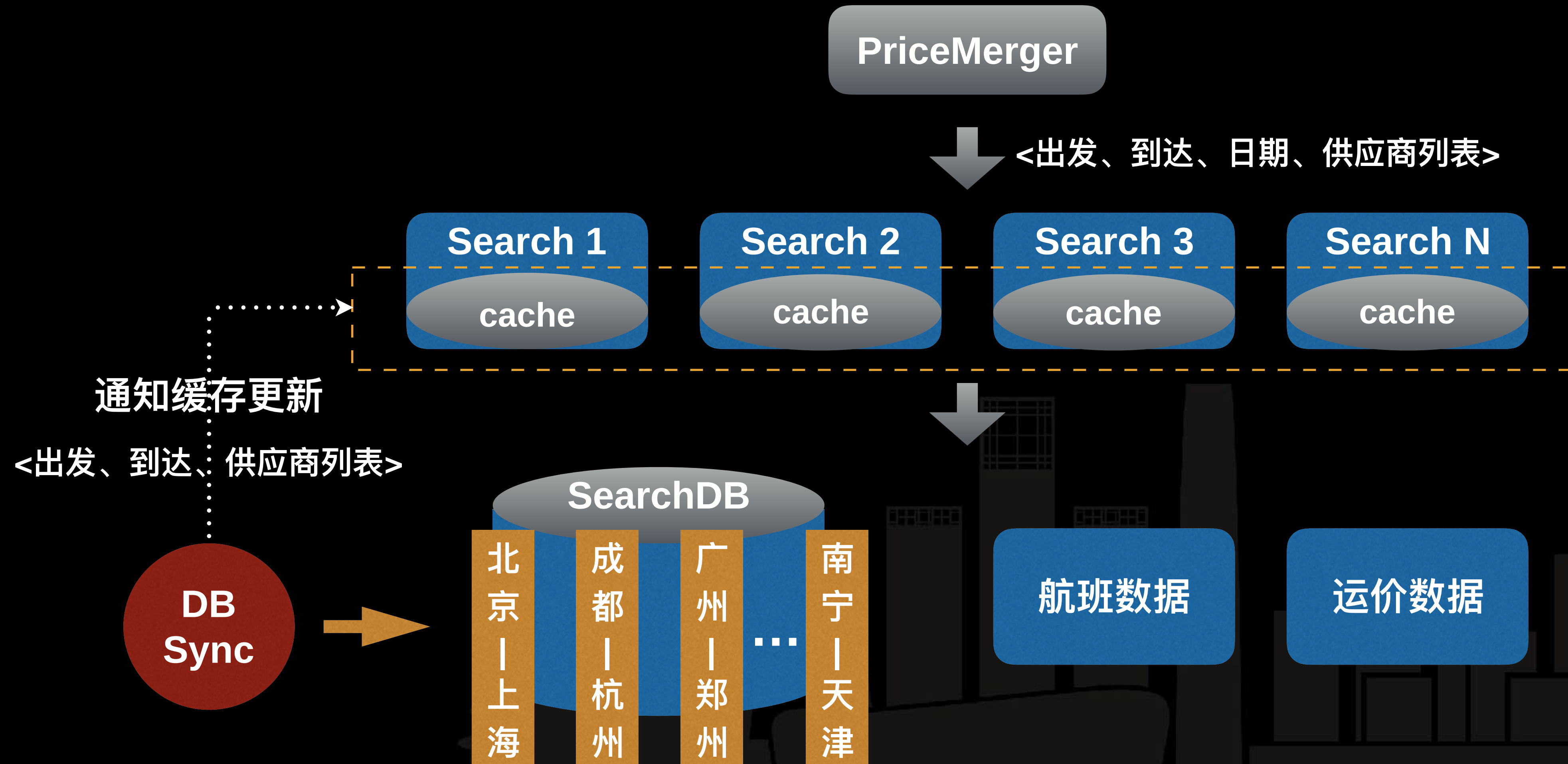


报价引擎 — 负载均衡

- 一致性哈希
- 缓存命中率
- 热点航线均衡
- 系统高可用

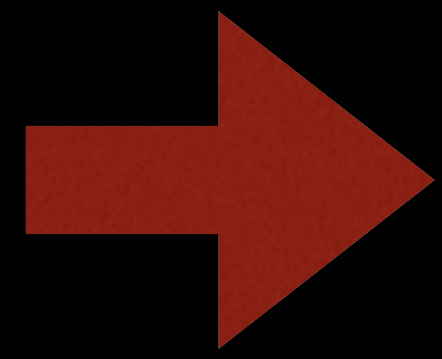


报价引擎 — 架构



报价引擎 — 小内存

- String
- ArrayList
- Range
- Date
- Integer
- Set



- byte, byte[]
- short, short[]
- int, int[]
- obj pool

报价引擎 — 性能优化

- 异步、并行、无锁化
- 剪枝
- 空间换时间
- 缩短对象驻留内存时间，减少gc次数，优化单机吞吐
- 数据交换采用protobuf + gzip处理
- jit、预热

回顾

- 水平分层，纵向分渠道，良好的扩展性
- 实时计算 + 阶梯式缓存，成本与报价新鲜度的权衡
- 闭环系统
- 索引库数据同步
- 本地缓存的设计，更新策略
- 缩减对象内存
- 一致性哈希负载均衡

心得

- 不同的业务场景，不同的特征
- 结合特征去进行设计和优化
- 通用!=最优
- 量体裁衣



待解决问题

- 开舱发现
- 国际运价计算
- ...





扫一扫上面的二维码图案，加我微信

Thanks!



主办方 **Geekbang** > **InfoQ**
极客邦科技